COMPUTER SCIENCE TRIPOS Part IB

Wednesday 7 June 2000 1.30 to 4.30

Paper 5

Answer **five** *questions.*
No more than **two** *questions from any one section are to be answered.*
Submit the answers in five **separate** *bundles, each with its own cover sheet. On each cover sheet, write the numbers of* **all** *attempted questions, and circle the number of the question attached.*
Write on **one** *side of the paper only.*

SECTION A

1 Data Structures and Algorithms

Explain what is meant by the terms *directed graph*, *undirected graph* and *bipartite graph*. [3 marks]

Given a bipartite graph, what is meant by a *matching*, and what is an *augmenting* path with respect to a matching? [4 marks]

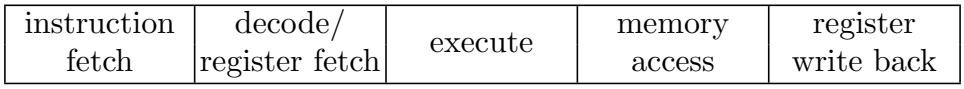
Prove that if no augmenting path exists for a given matching then that matching is maximal. [6 marks]

Outline an algorithm based on this property to find a maximal matching, and estimate its cost in terms of the number of vertices n and edges e of the given bipartite graph. [7 marks]

2 Computer Design

Modern processors exploit instruction level parallelism in order to improve throughput.

(*a*) Control-flow processors use pipelining to improve throughput. What limits throughput in the simple 5-stage pipeline depicted below? [10 marks]



(*b*) What is the principle of operation of a static data-flow processor and how does it resolve data dependencies? [10 marks]

3 Digital Communication I

What is meant by the term *flow control*? [3 marks]

What is meant by the term *credit-based* flow control? [4 marks]

What is meant by *start–stop* (or XON–XOFF) flow control? [4 marks]

A start–stop system is used on a 10 kbps link with a constant delay of 5 ms. How much buffer must a receiver keep in reserve for "stopping time" in order to prevent information loss? [3 marks]

Which system is more appropriate to use across the Internet and why? [6 marks]

4 Computer Graphics and Image Processing

Explain how a cathode ray tube (CRT) works, including details of how colour is achieved. [8 marks]

Describe a run-length encoding scheme for encoding images whose pixels have eight-bit intensity values. [8 marks]

Calculate the *best* possible compression ratio achievable with your scheme and describe the situation(s) in which this ratio would be achieved. [2 marks]

Calculate the *worst* possible compression ratio achievable with your scheme and describe the situation(s) in which this ratio would be achieved. [2 marks]

SECTION B

5 Comparative Programming Languages

Outline how you would implement complex numbers in C++. Your implementation should attempt to make complex numbers look as if they were built into the language by allowing new complex numbers to be declared, initialised, assigned and operated on by the normal arithmetic operators. [13 marks]

Discuss to what extent a good C++ compiler could implement your version of complex numbers as efficiently as if they had been a primitive type in the language. [7 marks]

6 Compiler Construction

Give a brief description of the main features of *either* Lex and Yacc *or* the corresponding Java tools JLex and Cup. [5 + 5 marks]

Illustrate their use by outlining how you would construct a parser for expressions composed of identifiers, integers, unary minus and binary operators $+$, $-$, $*$ and $/$. Your parser is expected to create a parse tree in a format of your choice representing the expression that is presented to it. If it helps, you may assume that expressions will be terminated by writing a semicolon after them. [10 marks]

7 Prolog for Artificial Intelligence

Consider the following problem to be solved using a Prolog program:

Given a closed planar polygon chain represented as a list of n vertices

$$\texttt{[v(}x_1\texttt{,}y_1\texttt{), v(}x_2\texttt{,}y_2\texttt{), ..., v(}x_n\texttt{,}y_n\texttt{)]}$$

compute the area of the enclosed polygon, and the orientation of the chain. The area is computed by the line integral $1/2 \int x \, \mathrm{d}y - y \, \mathrm{d}x$ where the integral is over the polygon chain. A naïve solution is given by the following program, which defines the predicate `area`. The goal `area(Chain,Area)` succeeds when `Chain` is the list of vertices, and the magnitude of `Area` is the area of the polygon bounded by the chain. The sign of `Area` is positive if the orientation of the polygon is anticlockwise and negative if it is clockwise:

```
area([X],0).
area([v(X1,Y1),v(X2,Y2)|VS],Area):-
          area([v(X2,Y2)|VS],Temp),
          Area is Temp + (X1 * Y2 - Y1 * X2) / 2.
```

Explain how vertices are processed by this procedure. [4 marks]

Why does this program execute inefficiently? [3 marks]

Write an alternative definition that is tail-recursive and makes use of accumulator variables. [10 marks]

Explain why your alternative definition executes more efficiently. [3 marks]

8 Databases

What is meant by a *functional dependency* between sets of attributes in a relational database schema? What conditions must be satisfied for a relation to be in Boyce–Codd Normal Form (BCNF)? [4 marks]

The Department of Transport is implementing plans to tax traffic congestion. From 2002, cars will carry approved radio-control units which at first will be used only to monitor vehicle movement. In controlled areas sensors identify all vehicles, recording their positions periodically. Amber signs flash when overall traffic flow drops below some threshold, and vehicles within the controlled area may be fined for lack of progress.

The owner of each vehicle has an account with the Department of Transport; owners can transfer funds to ensure that their account is in credit. Once credit is exhausted the level of fine increases by a factor 3, and a summons is sent by mail to the vehicle owner's registered address. In order to maintain proper accounts it is essential to keep an accurate record of each monitored offence.

You are employed to design the relational database that will enforce the scheme, including provision for vehicle and driver registration, monitoring of vehicle offences and management of vehicle accounts. Describe the schema you propose, stating clearly any assumptions that you make. You need not discuss the calculation of the fines due. [12 marks]

Outline the flow of information through the database. To what extent does the application require real-time transaction programming? [4 marks]

SECTION C

9 Semantics of Programming Languages

The *pish* shell features commands which return an integer exit status. Write

$$(*) \qquad C, s \Downarrow n, s'$$

to indicate that command C started in state s terminates with exit status n and final state s'. An exit status of zero indicates *normal* termination and a non-zero status indicates *abnormal* termination. The different forms of *pish* command are:

skip (immediate normal termination)

$C\,;C'$ (execute C and if it terminates normally continue with C'),

if B **then** C **else** C' (execute C or C' according to the value of the boolean expression B)

if B **return** n (return exit status n if B is true, otherwise terminate normally)

C **handle** n **with** C' (execute C and if it terminates with status n, execute C')

In all cases the final exit status and state is that produced by the last command executed. Define the structural operational semantics of these commands by giving an inductive definition of $(*)$. You may assume there is a relation of the form $B, s \Downarrow \mathbf{b}$ (where $\mathbf{b} \in \{\mathbf{true}, \mathbf{false}\}$) which defines the value of each boolean expression B in state s. [7 marks]

Write $C \cong C'$ to mean that for all s, n and s', it is the case that $C, s \Downarrow n, s'$ holds if and only if $C', s \Downarrow n, s'$ does. Show how to construct *pish* commands C_1, C_2 and C_3 from C, C', B and **true** just using the "**if** $-$ **return** $-$" and "$-$ **handle** $-$ **with** $-$" constructs so that

(a) $C_1 \cong$ **skip** [2 marks]

(b) $C_2 \cong C\,;C'$ [4 marks]

(c) $C_3 \cong$ **if** B **then** C **else** C' [7 marks]

Justify your answer in each case.

10 Foundations of Functional Programming

Give a brief account of how *four* of the following features of general programming systems can be modelled in terms of a form of un-typed functional programming where none of the mentioned facilities are provided as built-in features.

When selecting your examples and preparing your explanations, arrange that at least *one* of the four cases could be carried out using a typical polymorphically typed functional language while at least *one* would lead to type-checking problems.

(a) Tuples (it will be sufficient to consider just the case of pairs).

(b) Boolean quantities and an *if*/*then*/*else* construct.

(c) Lists (both empty and non-empty).

(d) Recursive function definitions.

(e) The numbers $0, 1, 2, \ldots$, with the associated operations of a zero test, addition and multiplication.

[4 marks each]

Explain the issues about type checking for all of the examples you have given.
[4 marks]

11 Logic and Proof

Given a propositional formula, we wish to test whether it is a tautology and, if it is not, to compute an interpretation that makes it false. Two techniques for doing this are the sequent calculus and ordered-binary decision diagrams. Give a brief outline of these techniques, applying both of them to the formulae

$$(A \to B) \to (B \to A) \quad \text{and} \quad (A \vee B) \to (\neg B \to A)$$

[7 + 7 marks]

It is proposed to replace the usual sequent calculus rule for disjunction on the left by this rule:

$$\frac{\Gamma, A \Rightarrow \Delta \quad \Gamma, B \Rightarrow \Delta, A}{\Gamma, A \vee B \Rightarrow \Delta}$$

Is this rule sound? Justify your answer. [3 marks]

Give an example to show that using this rule instead of the usual one makes some proofs shorter. [3 marks]

12 Complexity Theory

Give precise definitions of *polynomial time reductions* and *NP-completeness.* [2 marks each]

Consider the following two decision problems on *undirected* graphs.

> **3-node-colourability**: the collection of graphs $G = (V, E)$ for which there is a mapping $\chi : V \to \{r, g, b\}$ such that if $(u, v) \in E$, then $\chi(u) \neq \chi(v)$.
>
> **3-edge-colourability**: the collection of graphs $G = (V, E)$ for which there is a mapping $\chi : E \to \{r, g, b\}$ such that if $(u, v), (u, v') \in E$, with $v \neq v'$, then $\chi(u, v) \neq \chi(u, v')$.

Show that there is a polynomial time reduction from **3-edge-colourability** to **3-node-colourability**. [8 marks]

The problem **3-edge-colourability** is known to be NP-complete. Using this information, for *each* of the following statements, state whether or not it is true. In each case, give complete justification for your answer.

(a) There is a polynomial time reduction from **3-node-colourability** to **3-edge-colourability**. [3 marks]

(b) **3-node-colourability** is NP-complete. [3 marks]

(c) **3-edge-colourability** is in PSPACE. [2 marks]

END OF PAPER